Using low-resolution SAR scattering features for ship classification

Emanuele Salerno

National Research Council of Italy
Institute of Information Science and Technologies, Pisa, Italy

ABSTRACT

This letter reports an experimental study aimed at establishing the questionable usefulness of scattering attributes for ship classification from moderate-resolution SAR images. About 2700 example images representing four ship types have been extracted from the OpenSARShip annotated data set and used to form the training and test sets for random forest models. After importance ranking and cross-validation, different subsets of both geometric and scattering attributes were selected from a fixed training set and used to train the classifier. The results from the validation using the test sets show that the scattering attributes give a significant contribution in terms of overall classification accuracy.

***Index Terms*—** Low-resolution SAR ship classification; Scattering attributes; Random Forests

1. INTRODUCTION

Marine traffic surveillance is becoming more and more important as the increase in the number of vessels makes it essential to face different issues, such as border control, fisheries regulations enforcement and monitoring of illegal activities, as well as for general security and emergency management. To date, this is made possible on large scale by terrestrial or space-based identification systems, such as AIS and LRIT, and other systems deployed by the national maritime authorities.[1] In areas not covered by vessel traffic services, or when dealing with vessels that are non-collaborative or send falsified messages, ship detection and identification through remote-sensed images remains the only possibility to accomplish the above-mentioned missions. Satellite-borne SAR systems are particularly suited to this task for their insensitivity to weather and light conditions and their global coverage in very large swaths. SAR images offer the possibility of using Doppler processing and polarimetry to discriminate between different types of target, but their appearance is different from the usual optical images for a marked dependence on the imaging geometry and spatial resolutions of meters to tens of meters. This letter deals with resolutions of tens of meters, often referred to as moderate-resolutions in the literature, such as the ones provided by the C-band SAR system onboard the Sentinel-1 satellites,[2] and uses random forests [1], a non-deep-learning classification method, to identify different types of ships. The choice of moderate-resolution images makes sense, as a high-resolution SAR image is obtained at the cost of a narrow swath, which is unsuitable for maritime surveillance [2]. Despite deep learning is proving to be very effective in marine applications [3, 4], the results produced from moderate-resolution data are not so overwhelmingly better than those obtained by classical algorithms using hand-crafted features [5, 6].

With moderate-resolution images, the non-deep-learning algorithms often perform comparably to the deep-learning ones [7]. Many authors, however, maintain that the only significant features to be used are the geometric ones [2, 8, 9]. In [2], it is suggested that with such images the so-called 'naive' geometric features (NGF), only depending on length and width of the target vessel, are sufficient for classification. Indeed, the spatial distribution of the strong scattering centers in moderate-resolution images has not the clear patterns shown by high-resolution images, thus their discriminative power is limited. The results presented here contradict this claim, as we always found some scattering features in the subsets that produce the best classification accuracies. We selected four ship types to be classified – bulk carriers, cargo ships, container ships and tankers – and extracted a number of annotated images thereof from the OpenSARShip data set [10], featuring Sentinel-1 wide-swath, ground-range-detected images (IW-GRDH), with a spatial resolution of $20\mathrm{m} \times 20\mathrm{m}$ and a pixel size of $10\mathrm{m} \times 10\mathrm{m}$. By 2021, these four types constitute more than an 88% of the global merchant fleet [11]. Following [2], we chose the NGFs as geometric attributes and, inspired by other suggestions found in the literature [8, 12], we chose electromagnetic features

This work has been partly supported by ESA GSTP Project OSIRIS-FO (Optical/SAR data and system Integration for Rush Identification of Ship models – Follow-On), contract 4000132250/20/I-NB.

[1] https://www.imo.org/en/OurWork/Safety/Pages/AIS.aspx; https://www.imo.org/en/OurWork/Safety/Pages/LRIT.aspx; https://www.imo.org/en/OurWork/Safety/Pages/VesselTrafficServices.aspx

[2] https://sentinel.esa.int/web/sentinel/missions/sentinel-1

This article has been accepted for publication in IEEE Geoscience and Remote Sensing Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LGRS.2022.3183622

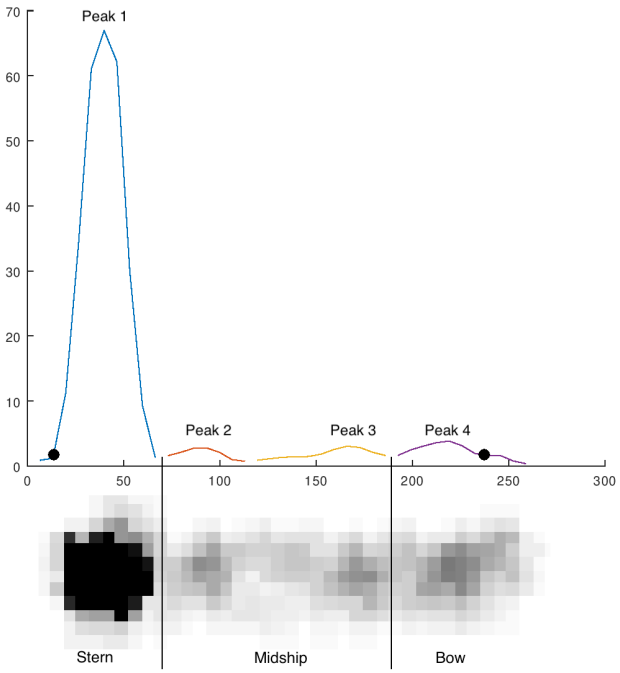


Fig. 1. Top: Longitudinal scattering profile from a typical ship footprint. Horizontal scale in meters, vertical scale in absolute, calibrated scattering values. Bottom: matched target image divided into the stern, midship and bow sections.

Table 1. Geometric and scattering features considered.

Feature	**Symbol/Formula**	**Notes**
Length	L	Estimated
Width	W	”
Perimeter	$P = 2(L + W)$	Naive def. [2]
Area	$A = L \cdot W$	”
Elongatedness	$El = L/W$	”
Aspect ratio	$AR = W/L$	”
Circularity	$Ci = 4LW/(L + W)^2$	”
Compactness	$Co = (L + W)/L$	”
Mean, Overall	avg_all	
No. of Peaks	N_pks	In the mean scattering profile
Pos. of Max Peak	l_mxpk	From stern
Value, Max Peak	v_mxpk	
No. of Top Peaks	N_tpks	Higher than $0.5 \cdot v\_mxpk$
Mean, Stern	avg_strn	From the mean scattering profile
Mean, Midship	avg_mid	”
Mean, Bow	avg_bow	”

related to the mean scattering values in different sections of the ship footprint and the distribution of the peaks in the scattering profile. Using the images and the ground truth provided by OpenSARShip, we trained random forest models with different subsets of the geometric and scattering features and compared the results obtained using only geometric, only scattering and both geometric and scattering features. It emerged that significantly improved classification performances are obtained using both geometric and scattering features.

2. FEATURES AND DATA SET

The algorithm we used to extract the ship footprint from each SAR image is described in [13], with a refinement inspired by [14]. The ship length and width are estimated from this refined footprint as its maximum lengths along the principal inertia axes. By averaging the scattering values along lines orthogonal to the keel (that is, orthogonally to the minimum inertia axis), we obtain a scattering profile normally characterized by several peaks (Fig. 1). The highest between the first and the last peaks detected is referred to as ‘Peak 1’; all the other peaks are labeled by progressive numbers. Since normally the stern section scatters more than the bow, we assume that the stern corresponds to our Peak 1. Then, we locate the boundaries of the stern, midship and bow sections in the profile relative minima that are the closest, respectively, to one third and two thirds of the ship length. The first four of our scattering features are the mean scattering values over the whole ship footprint and the three sections thus defined. The remaining features are the label and the value of the highest peak in the profile, the total number of peaks and the number of peaks that exceed half the maximum. We thus have eight scattering features that, along with the eight NGFs, form our overall feature set, described in Table 1.

Each Sentinel-1 scene in OpenSARShip is accompanied by an annotation xml file with the ground-truth of each target and a folder containing two-channel tiffs with the vv- and vh-polarized, calibrated image chips representing all the listed targets. From the 2755 image pairs extracted, we first eliminated the ones with no ground-truth size available and then applied our feature extraction algorithm. The estimated geometric features were often affected by very large errors caused by the different clutter levels, imaging artifacts and other anomalies in the data. Removing the targets with relative errors larger than a 40%, however, still leaves a sufficient number of records to be processed. In any case, the estimation algorithm still offers the possibility to be refined and corrected for possible systematic errors. Finally, 1418 vh-polarized images with 39 ship types and 1332 vv-polarized images with 37 ship types survived this selection. From the

vh targets, we extracted 521 bulk carriers, 443 cargo ships and 164 container ships. Then, assuming that their features are similar, we grouped together in the type "Tanker" the 122 ships marked as "Bunkering Tanker", "Chemical Tanker", "Crude Oil Tanker", "Oil Products Tanker", "Oil/Chemical Tanker", or "Tanker". The same process on the vv images produced 515 Bulk Carriers, 406 Cargo, 167 Container Ships and 100 Tankers. The features extracted from the vh data set were used to build a perfectly balanced training set with 100 randomly chosen ships per type. From the remaining 850 records, we built a test set with 100 bulk carriers, 100 cargo, 64 container ships and 22 tankers. The no-information rate (NIR) for such a set is 0.35. A classification task can be considered successful when the resulting accuracy lays above the NIR. For the vv images, with the same modalities, we vere able to build a training set with 80 ships per type and a test set with 20 ships per type, with a NIR of 0.25. The results in the vv case would probably be less reliable than in the vh case because of the smaller training and test sets used.

3. EXPERIMENTS

To train and test the random forest models, we used the R packages `randomForest` and `caret`, which include useful functions to produce and evaluate the results.[3] Through these functions, we are able to compute the relative feature importances, train random forest models using different feature subsets and then evaluate their performances using the test sets. The models have been trained through the feature values as estimated by the algorithm, with no centering or normalization. This choice was motivated by the need to leverage the actual sizes and scattering amplitudes to characterize the ship types. This is also why we used the floating-point calibrated image chips rather than their 16-bit unsigned integer versions. A series of experiments has been tried using the principal components of the normalized features. The results obtained are not reported here, as they showed no relevant difference with respect to the ones presented below.

The importance ranking for the geometric features in the vh case is as follows (values in parentheses): 1) P (50.96); 2) L (48.16); 3) A (47.22); 4) W (36.15); 5) Ci (26.21); 6) Co (25.70); 7) AR (25.49); 8) El (25.22). From a 5-fold cross validation computed through function `rfcv`, the minimum relative classification error, 0.535, is obtained by using the first 3 features. To validate this result through never-seen-before data, we trained 7 models, with 2 to 8 features, and tested them against our 286-record test set. The results, in terms of the overall accuracies and the related 95% confidence intervals as functions of the number of features, are reported in Fig. 2(a). The maximum accuracy, 0.465, is obtained with just two features, P and L.

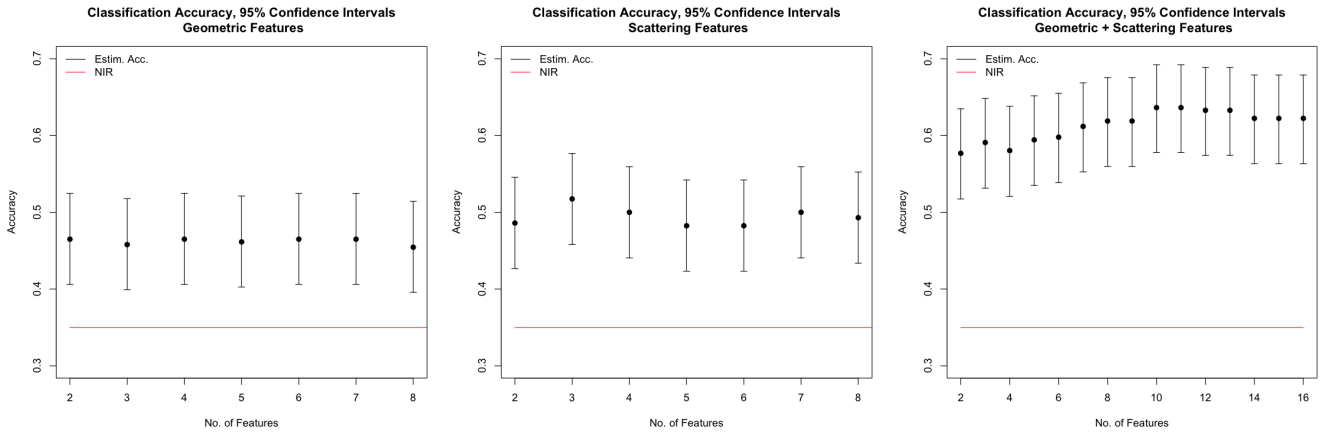
The same procedure has been followed for the scattering features; the result is summarized in Fig. 2(b). The importance ranking is in this case 1) *avg_mid* (57.01); 2) *v_mxpk* (51.55); 3) *avg_bow* (50.73); 4) *avg_all* (47.93); 5) *avg_stern* (46.88); 6) *N_pks* (16.28); 7) *l_mxpk* (13.93); 8) *N_top.pks* (10.90). The minimum error obtained by cross validation is 0.502, with all the 8 features, whereas the maximum accuracy using the test set, 0.518, is obtained with three features: *avg_mid, v_mxpk* and *avg_bow*. Note that the accuracy reached through scattering features alone is comparable to the one obtained through geometric features alone. Indeed, all the confidence intervals obtained overlap in the range [0.458, 0.514]. Considering that these intervals are entirely above the NIR; this means that our scattering features are indeed significant for the classification of these four types of ships.

The importance ranking of all the 16 features is 1) *avg_mid* (34.13); 2) P (33.07); 3) L (28.89); 4) A (27.97); 5) *avg_bow* (27.19); 6) *avg_all* (23.23); 7) W (20.42); 8) *avg_stern* (20.35); 9) *v_mxpk* (20.11); 10) Ci (11.28); 11) Co (11.28); 12) El (10.87); 13) AR (10.64); 14) *l_mxpk* (8.47); 15) *N_pks* (5.93); 16) *N_top.pks* (5.40). The minimum error by cross validation, 0.385, is obtained with 11 features. Note that the top ranking feature is of scattering type, confirming that these are useful with moderate-resolution images. Among the 11 most important features, there are 5 scattering features, namely, the ones directly related to the scattering values. The remaining three, only related to number and position of the peaks in the scattering profile, are at the bottom of the ranking. The results of the accuracy evaluation using the test set are summarized in Fig. 2(c). The maximum accuracy, 0.636, is obtained with 10 and 11 features. Also, a simple comparison of the three plots in Fig. 2 shows that not only the scattering features are as discriminative as the geometric ones for the four ship types considered, but also that the use of both kinds of features produces an apparent advantage in terms of overall classification accuracy.

The analysis of the vv-polarization case confirmed the usefulness of the scattering features for classification. Since our fundamental claim is thus demonstrated, here we only report the results from the cross validation, showing that the classification errors obtained using both types of features are significantly smaller than the ones obtained through either the NGFs or the scattering features alone. Due to the smaller training and test sets, in this case the confidence intervals around the accuracy values are wider than in the vh-polarization case. The importance ranking of the features extracted from the vv images is 1) P (29.01); 2) L (26.06); 3) A (22.52); 4) *avg_mid* (19.00); 5) *avg_bow* (16.85); 6) W (16.45); 7) *avg_stern* (15.77); 8) *avg_all* (15.47); 9) *v_mxpk* (13.78); 10) *l_mxpk* (10.97); 11) AR (10.96); 12) Ci (10.87); 13) El (10.60); 14) Co (10.02); 15) *N_pks* (6.25); 16) *N_top.pks* (4.65). The usual 5-fold cross validation found a minimum error of 0.434 using 15 features.

[3] `https://www.rdocumentation.org/packages/randomForest/versions/4.6-14/topics/randomForest`; `https://www.rdocumentation.org/packages/caret/versions/6.0-91`

This article has been accepted for publication in IEEE Geoscience and Remote Sensing Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LGRS.2022.3183622



(a) (b) (c)

Fig. 2. Classification accuracies and 95% confidence intervals, vh polarization, as functions of the number of features: (a) geometric features alone; (b) scattering features alone; (c) both geometric and scattering features.

The above-reported ranking confirms, for vv-polarized images too, that at these resolutions the features related to the distribution of strong scattering points, namely, *l_mxpk*, *N_pks* and *N_top.pks*, are not very important for classification. Conversely, the scattering features directly related to scattering amplitudes are always among the most important ones, and contribute significantly to increase the accuracy. Indeed, the cross validation finds a minimum error of 0.547, with 5 features, using the NGFs alone, and a minimum error of 0.597 using all the 8 scattering features. It can be noted that the error obtained with 15 features is much smaller than the last two figures. The accuracy in classifying the test set, using the 15 most important features of both types, is 55%, with 95% confidence interval [0.43, 0.66]. Using the 5 most important NGFs, we obtain an accuracy of 47.5% with confidence interval [0.36, 0.59]. Finally, using all the scattering features we obtain an accuracy of 48.8% with confidence interval [0.37, 0.60].

4. DISCUSSION

This experimental work demonstrates how some scattering features offer a significantly beneficial effect in ship classification from even moderate-resolution SAR images. In the case treated here, the best classification has been obtained from vh-polarized images using a feature subset with 11 entries. Richer training sets would be needed to fully exploit the discriminative potentialities of the whole feature set. Strictly, the results obtained with vv-polarized images could not be compared to the vh ones, as the vv-based training set was less populated than the other. A further study will use more data, thus allowing for a fair comparison.

Even though trying to maximize the classification performance was out of our scope, we observe that the accuracies reported here are comparable to other results presented in the literature regarding non-deep-learning classification models applied to moderate-resolution images [15]. Interestingly, in [16], a CNN is used for classification, also leveraging the incident angle extracted from the SAR metadata in view of the fact that the appearance of a SAR target depends on the imaging geometry. Despite the claimed improvement over a CNN with no SAR-geometry integration, the results for the three classes Container, Bulk Carrier and Tanker produce a maximum accuracy of 74.5% using a total of 1371 ship chips from OpenSARShip. Other deep learning strategies proposed claim results that are generally much better than the ones mentioned so far, with the advantage of avoiding the use of handcrafted features. In [3, 4], accuracies close to 100% are obtained by using pretrained convolutional networks with, respectively, fine-tuning and transfer learning from small sets of high-resolution SAR images. The four CNN models used in [5] distinguish among cargo ships, tankers, platforms, windmills, and harbor structures from high-resolution TerraSAR-X data, with F1-scores from 59% to 80% for cargo and tankers and close to 100% for the other structures. A fifth CNN model, trained with multiresolution data, produces F1-scores of 85% and 87% for cargo and tankers, respectively. In [17], a custom-made CNN is trained through variously augmented training sets from few originals and obtains very good results on 8 types of marine objects imaged at resolutions between 0.5m to 5m. Again, despite the high resolution in input and although the accuracies reached are generally very high, the figures obtained for ships are smaller than the ones obtained for platforms and other classes of marine targets. Maybe hybrid approaches, such as the ones proposed in [18], complementing handcrafted features and a deep learning approach, would deserve to be examined in depth in the future.

5. CONCLUSION

The aim of this letter was to demonstrate experimentally that using both geometric and scattering attributes extracted from moderate-resolution SAR images is advantageous for ship classification over using only one kind of attributes. The results shown confirm this conjecture. The data sets used are not so large to allow all the features devised to be included in the training and no optimization has been attempted in the feature extraction and classification algorithm. This notwithstanding, the classification performances obtained are comparable to several results reported in the literature.

6. REFERENCES

[1] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[2] H. Lang and S. Wu, "Ship classification in moderate-resolution SAR image by naive geometric features-combined multiple kernel learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1765–1769, 2017, doi: 10.1109/LGRS.2017.2734889.

[3] Y. Wang, C. Wang, and H. Zhang, "Ship classification in high-resolution SAR images using deep learning of small datasets," *Sensors*, vol. 18, no. 2929, 2018, doi: 10.3390/s18092929.

[4] C. Lu and W. Li, "Ship classification in high-resolution SAR images via transfer learning with small training dataset," *Sensors*, vol. 19, no. 63, 2019, doi: 10.3390/s19010063.

[5] C. Bentes, D. Velotto, and B. Tings, "Ship classification in TerraSAR-X images with convolutional neural networks," *IEEE Journal of Oceanic Engineering*, vol. 43, no. 1, pp. 258–266, 2018, doi: 10.1109/JOE.2017.2767106.

[6] T. Zhang, X. Zhang, X. Ke, C. Liu, X. Xu, X. Zhan, C. Wang, I. Ahmad, Y. Zhou, D. Pan, J. Li, H. Su, J. Shi, and S. Wei, "HOG-ShipCLSNet: A novel deep learning network with HOG feature fusion for SAR ship classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–22, 2022, doi: 10.1109/TGRS.2021.3082759.

[7] G. Zhou, G. Zhang, and B. Xue, "A maximum-information-minimum-redundancy-based feature fusion framework for ship classification in moderate-resolution SAR image," *Sensors*, vol. 21, no. 2, p. 519, 2021, doi: 10.3390/s21020519.

[8] H. Lang, J. Zhang, X. Zhang, and J. Meng, "Ship classification in SAR image by joint feature and classifier selection," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 2, pp. 212–216, 2016, doi: 10.1109/LGRS.2015.2506570.

[9] B. Snapir, T. W. Waine, and L. Biermann, "Maritime vessel classification to monitor fisheries with SAR: Demonstration in the north sea," *Remote Sensing*, vol. 11, no. 353, 2019, doi: 10.3390/rs11030353.

[10] L. Huang, B. Liu, B. Li, W. Guo, W. Yu, Z. Zhang, and W. Yu, "OpenSARShip: A dataset dedicated to Sentinel-1 ship interpretation," *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, vol. 11, no. 1, pp. 195–208, 2018, doi: 10.1109/JSTARS.2017.2755672.

[11] United Nations Conference on Trade and Development. Data center. Last checked March 18, 2022. [Online]. Available: https://unctadstat.unctad.org/wds/ReportFolders/reportFolders.aspx

[12] M. Jiang, X. Yang, Z. Dong, S. Fang, and J. Meng, "Ship classification based on superstructure scattering features in SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 616–620, 2016, doi: 10.1109/LGRS.2016.2514482.

[13] L. Bedini, M. Righi, and E. Salerno, "Size and heading of SAR-detected ships through the inertia tensor," *Proceedings (MDPI)*, vol. 2, no. 97, 2018, doi: 10.3390/proceedings2020097.

[14] M. Stasolla and H. Greidanus, "The exploitation of Sentinel-1 images for vessel size estimation," *Remote Sensing Letters*, vol. 7, no. 12, pp. 1219–1228, 2016, doi: 10.1080/2150704X.2016.1226522.

[15] A. Makedonas, C. Theoharatos, V. Tsagaris, and S. Costicoglou, "A multilevel approach to ship classification on Sentinel-1 SAR images using artificial neural networks," in *LPS16 ESA Symposium*, 2016.

[16] S. Sharma, K. Senzaki, and H. Aoki, "CNN-based ship classification method incorporating SAR geometry information," *SPIE Proceedings*, vol. 10789, 2018, doi: 10.1117/12.2325282.

[17] M. Ma, J. Chen, W. Liu, and W. Yang, "Ship classification and detection based on CNN using GF-3 SAR images," *Remote Sensing*, vol. 10, no. 2043, 2018, doi: 10.3390/rs10122043.

[18] T. Zhang and X. Zhang, "Injection of traditional hand-crafted features into modern CNN-based models for SAR ship classification: What, why, where, and how," *Remote Sensing*, vol. 13, no. 11, p. 2091, 2021, doi: 10.3390/rs13112091.